

Haile Quals day 4

Some of your readings cover the development of expertise (or skill) from a cognitive perspective (mind rather than brain)

- *According to these readings, what are differences between experts and novices (those who are highly skilled and those in the initial stages of learning)?*
- *What changes as a person becomes more skilled or expert in a particular domain?*
- *Are there differences/disputes/disagreements with respect to what is taking place in the development of expertise (or skill) among the views represented in your readings? If so, explain them and offer your own opinion on the topic.*
- *Finally, what are the implications of this issue for our understanding of human cognition in general?*

All skill learning theories bear the same hallmarks of expert performance. Experts can execute tasks efficiently with little mental effort and at much higher levels of accuracy than novices. In most theories, this is a multi-stage process but the number of stages and the reason for the improvement tend to vary or are non-specific. Current theories seem to have been influenced by the much earlier Fitts and Posner (1967), Schneider and Shiffrin (1977), and Anderson, (1982) models so these models are discussed in this essay.

There are two big questions in skill learning: what does expert performance look like compared to novice performance? And how does this expert performance develop? These two questions are intimately linked because it seems like some of the models of skill development have started by observing skilled performance and tried to engineer a cognitive and psychological explanation for how they came about.

There are many terms in psychology that describe a minimum number of 2 psychological states concerned with task performance and these seem to rely on the features of different stages of skill performance. Expert performance is always described as being fast, sparing of cognitive resources, effortless, virtually error free, automatic, and relatively inflexible. This performance is almost habitual. Novice learners are often characterized by their effortful, slow, and error-prone performance. Novices have to pay attention to all aspects of the environment and their response patterns (Fitts and Posner, 1967; Schneider and Shiffrin, 1977; Anderson, 1982; Tenison et al., 2016).

With proponents of 2-system theories, early-stage performance is termed *controlled processing* and late-stage, or expert performance is known as *automatic processing* (Schneider and Shiffrin, 1977; Hill and Schneider, 2006). These theories are influenced by early visual attention and visual search experiments. Schneider and Shiffrin's 2-system theory was demonstrated by visual search tasks where subjects were required to find and report items from a specific category and ignore other categories of objects. In varied mapping experiments, where

the target category changed trial-by-trial, subjects had to rely on serial search methods that required attentional engagement and there was not much improvement in speed (controlled processing). But there was an improvement in speed and accuracy with subjects who experienced consistently mapped visual search experiments, where the target category does not change from one trial to the other, because they were able to engage in automatic processing. Therefore, controlled processing must be engaged in new and constantly changing situations (Hill and Schneider, 2006). Additionally, controlled processes are affected by task load and fatigue whereas automatic processing is largely not, pointing to differential use of available cognitive resources.

For Schneider and Shiffrin, controlled processes had to rely on domain-general attentional and executive processing because the subject had to deploy attention to search stimuli in the environment and in long-term memory. But once sufficient training occurs in a reliable environment, this top-down, and effortful control and attention network is not needed. Automatic processes are represented in a non-specific long-term memory store where the representations (nodes, according to Schneider and Shiffrin) are activated automatically by current stimuli and responses are deployed without the subject directing them (Schneider and Shiffrin, 1977). The above-mentioned domain-general attentional resources are free to be utilized elsewhere. This notion has been exploited by various attention and skill learning studies using dual-task experiments to demonstrate that skilled performance does not rely on top-down attention (e.g., comparing expert golfers with novices while they memorized items from an auditory list - Beilock et al. 2002). This controlled and automatic processing has gone on to influence multiple lines of research, including more recent attempts to find more evidence for how this is implemented in the human brain through network neuroscience (see Hill and Schneider, 2006 for a review).

However, the 2-system theory described above does not well characterize how skill transitions from requiring a high level of control to automaticity, and the systems described seem non-specific. Other concurrent theories describe a multi-stage development of skill that are more specific in the types of long-term memory that guide novice and expert performance. The most notable of these are theories by Fitts and Posner (1967) and Anderson (1982). Both models are influenced by computation and computer structure theories typical to their time from the 1960 to the 1980s, that have prevailed in modern theories of cognition.

The Fitts and Posner (F&P) (1967) model of skill acquisition has broad strokes that are like the Schneider and Shiffrin model, but goes a little deeper with a hierarchically organized system that contains 3 learning stages: 1) Cognitive Phase, 2) Associative phase, and, finally, 3) Autonomous phase. The F&P model also tracks the familiar progression from effortful cognitive engagement in the early stages of learning to the more effortless and automatic late-stage learning. It should also be noted that progression through the stages occurs gradually and not in a staccato fashion even though they have forwarded distinct phases.

This model conceptualizes skill development as a streamlining process that constructs response 'routines' that are made-up of smaller, relevant, and flexible 'sub-routines' or habits. It

may be helpful to think of subroutines as the learner's store of stimulus-response maps and strategies and they seem very similar to Schneider and Shiffrin's long-term memory nodes. But the F&P model includes an additional hierarchical structure here that also harkens back to computing terminology, sub-routines. Sub routines are pieces of code or functions that perform repeated tasks and make up larger programs. Similarly in human cognition, F&P (1967) argue that the first cognitive phase is learning which subroutines are relevant, out of a multitude of available subroutines. This phase requires attending to and representing relevant features of stimuli, specific cues, and responses, which requires effort. During this stage, the learner normally applies inappropriate subroutines which lead to errors and poor performance. There is a reference to the learner's memory, and executive and attentional resources but not much else is described along the more influential descriptions of declarative and procedural memory/learning processes that come later in other models.

During the second stage, the associative phase, new subroutines are created, and new routines constructed from older subroutines are tested. Then a process of elimination, by attending to responses and cues, is performed where inappropriate subroutines that result in errors are tuned out or eliminated (this process of *tuning* will also appear in Anderson's skill learning theory, stay tuned). Successful routines are practiced and the number of activated inappropriate subroutines are reduced along with an appreciable reduction in errors.

The final autonomous phase occurs once the new routines are established and become automatic. There are virtually no mis-applied subroutines, and the learner no longer needs to deploy attentional and executive decision-making resources to parse available routines, and selects only those that led to success. There is virtually no explanation of how this occurs in the early text that I could find. There is however more explanation of behavioral features of decreases in errors and increases in speed and efficiency using this framework. Anderson (1982) goes a little further to establish a more detailed progression of skill learning that develops in more modern times.

Anderson's model of skill learning, while it shares some features with its early contemporaries, describes automatization of skill more specifically. It also forms the basis for the expansive, popular, and influential cognitive architecture, Adaptive Control of Thought-Rational (ACT-R) (Ye et al.,2018).

In Anderson's model (1982), all skills start out as declarative information also known as factual knowledge. Declarative instructions or rules stored in long-term memory, or newly acquired from the environment (e.g., reading instructions for a specific recipe from a cookbook), are attended to and the instructions are carried out laboriously. But factual knowledge in declarative memory only indirectly affects behavior. An additional procedural system interprets the instructions and carries them out. This procedural system contains simpler, domain-general *procedures* for carrying out common tasks, not unlike the Fitts and Posner (1967) subroutines. This is a process that is resource intensive and needs engagement with attention and verbal rehearsal of instructions that are effortfully retrieved from long-term memory. At least at this early stage, Anderson describes this as requiring space in working memory, a concept not

discussed in much detail by both the Schneider and Shiffrin and Fitts and Posner models. This early stage also aligns with observations of novice performance of slow, error prone work.

With practice, through a process of *knowledge compilation*, the learner will become less reliant on direct recall of instructions from declarative memory and laborious interpretation by the procedural system via working memory (Anderson, 1982). Compilation has two distinct sub-processes. The first, *composition*, describes the process of combining sequences of procedures into fewer procedures. The second, *proceduralization*, rebuilds these combined procedures in such a way that they no longer require retrieval of information from long-term memory. This is very similar to F&P (1967) association phase and results in speed-up of performance and reduction in use of working memory resources.

Anderson's skill learning theory has evolved over the years to seem more like human cognition than a computing framework. Knowledge compilation, in my opinion sounds very much like early computing concepts and terminology, though it goes a step further into explaining what occurs during skill development. This representation of two systems, the first that contains declarative knowledge and the second that learns *how* to act on that knowledge and eventually become independent of the first, much resembles how human cognition is organized. This is so at least from the point of view of numerous studies with amnesic patients. These studies have shown that damage to certain brain regions resulted in loss of some but not all memory functions. For example, patients with amnesia, can learn new procedural skills but not new declarative facts, or remember that they learned the skill (e.g., Packard and Knowlton, 2002; Anderson, 2002).

Anderson's cognitive architecture, ACT-R, has evolved through the years and was much informed by human and animal neuroscience studies (Anderson, 2007). ACT-R is modular and contains modules for sensory perception, attention, declarative memory, and procedural learning. But it also includes a system of goal representation. But I will discuss only the aspects of ACT-R that are relevant for skill learning, mainly in the procedural system.

Anderson's procedural system (Anderson, 1982; Anderson 2002) performs the bulk of the "learning" in skill learning. The functions of the procedural system largely explain how skills develop and their behavioral hallmarks.

Skilled performance resulting from practice is a gradual and slow process. It cannot be explained by the rapid learning that we would see in declarative systems. Therefore, this must occur slowly and with reward through reinforcement learning (Anderson, 2002). Reinforcement 'teaches' the learning system which procedures are useful. If a series of procedures are utilized to interpret declarative information and produce a behavior, those procedures are rewarded. This increases their chances of being utilized again given the requirements from the environment (Anderson, 2002). This implementation of reinforcement learning in the procedural system goes a long way to bridging the gap between model and a plausible biological implementation of learning, much further than the other discussed models.

Recent studies have used the ACT-R cognitive architecture to examine skill development in further detail. These studies have also overlaid ACT-R onto Fitts and Posner's (1967) three

phases cognitive, associative, and autonomous phase, resulting in improved characterization of where improvement occurs. For instance, Tenison et al. (2016) further divide the F&P phases into three stages: Encoding, Solving and Responding. They used an experiment where learners were given a specific mathematical problem with instructions on how to solve it. During the encoding stage, learners represent the problem in working memory. In the solving stage, learners retrieve the rules for solving the problem from declarative memory and apply it. And finally, they simply report the answer by typing out the computed numbers. Tenison et al. (2016) argue that the amount of time a learner spends in the encode, solve, and response stages change depending on which F&P phase — cognitive, associative, or autonomous — the learners are in. As learners practiced, they spent less time in the ‘solve’ stage because answers to the problems may be readily retrievable, and so, spend most of their time encoding and responding. Once this happens, they are more likely to be in associative phase than either of the cognitive or autonomous phases. Proceduralization can be so streamlined, with more practice, that retrieval of a response will not even be necessary as the response will be represented by motor patterns (Tenison et al., 2016; Anderson, 1982). This very rapid encoding of the problem and rapid, effortless response means that the subjects have reached an autonomous phase. Tenison et al., 2016, demonstrate this transition using behavioral, cognitive modeling and neuroimaging techniques.

The above study exemplifies the distance we have come in understanding skill learning. In summary, skill learning is a multi-stage, multi-component refining of responses or problem solving that is characterized by increase in accuracy and speed, and significant decrease in errors and cognitive effort. The discussed models and many others characterize how this occurs by attempting to explain how these measurable behavioral changes might occur. It seems like a lot of the early models were very influenced by computer architectures and computing concepts. They have forever altered how we think about cognition. We are likely to use computing terms like retrieval, storage, buffers, information processing, and compiling, mainly because of this legacy but they have allowed us to think more concretely about cognition as well.